Theoretical Computer Science 134 (1994) 403–413
Elsevier



# Representation of rational functions with prefix and suffix codings*

T. Harju

*Department of Mathematics, University of Turku, Yliopistonmäki, SF-20500, Turku, Finland*

H.C.M. Kleijn

*Department of Computer Science, Leiden University, P.O. Box 9512, 2300 RA Leiden, Netherlands*

M. Latteux and A. Terlutte

*Lab. d'Informatique Fondamentale de Lille, CNRS UA 369, Université de Lille I, F-59655 Villeneuve d'Ascq, France*

Communicated by M. Nivat
Received January 1993
Revised October 1993

*Abstract*

Harju, T., H.C.M. Kleijn, M. Latteux and A. Terlutte, Representation of rational functions with prefix and suffix codings, Theoretical Computer Science 134 (1994) 403–413.

We proceed with the characterization of rational functions by means of restricted class of morphisms. Left subsequential transductions can be factored in an endmarking followed by an uniform morphism, the inverse of a prefix morphism and an alphabetic morphism. Rational functions require the inverse of a prefix morphism followed by the inverse of a suffix morphism.

*Résumé*

Nous poursuivons l'étude des fonctions rationnelles et leurs caractérisations en termes de compositions de morphismes. Nous montrons que les transductions sous-séquentielles peuvent se factoriser en un marquage de fin de mot suivi d'un morphisme uniforme, d'un morphisme préfixe inverse et d'un morphisme alphabétique. Les fonctions rationnelles nécessitent l'emploi d'un morphisme préfixe inverse suivi d'un morphisme suffixe inverse.

*Correspondence to*: A. Terlutte, CNRS UA 369, Lab. d'Informatique Fondam. de Lille, Université de Lille I, Bât. M3, F-59655 Villeneuve d'Ascq, France.

* This work was partially supported by the Esprit Basic Research Action Working Group 6317 ASMICS 2, the PRC Mathématiques et Informatique and the Academy of Finland.


*SSDI* 0304-3975(93)E0199-E

## 1. Introduction

It was proved in [9, 12, 16] that a multivalued mapping $\tau$ is a rational transduction if and only if it can be factored in the form $\tau=\mu\sigma$, where $\mu$ is an endmarking and $\sigma=\alpha_1^{-1}\alpha_2\alpha_3^{-1}\alpha_4$ is a composition of morphisms and inverses of morphisms. In [7], a similar result was shown to hold for the rational functions. A mapping $\tau$ is a rational function if and only if it can be factored in the form $\tau=\mu\sigma$, where $\sigma$ is a composition of morphisms and inverses of injective morphisms and $\mu$ is an endmarking.

Further, in [8] the deterministic rational functions were specified in terms of compositions. It was proved there that a (partial) mapping $\tau$ is a deterministic rational function if and only if it has a factorization of the form $\tau=\alpha_1\mu\alpha_2^{-1}\rho\alpha_3$, where $\alpha_i$, $i=1,2,3$, are morphisms with $\alpha_2$ a prefix coding, $\mu$ is an endmarking and $\rho$ is a removal of endmarkers. Also, in [8], the subsequential rational functions were shown to consist of the compositions $\tau=\mu_m\alpha_1\alpha_2^{-1}\alpha_3$, where $\mu_m$ is a endmarker, and of the three morphisms, $\alpha_2$ is again a prefix coding.

In this paper we improve the results of [8] for a natural subclass of subsequential rational functions by showing (Theorem 3.6) that the positive and faithful subsequential functions require only restricted morphisms in the composition. Our main result (Theorem 4.6) states that if $\tau$ is a positive and faithful function, then

$$\tau=\mu\alpha_{\mathrm{u}}\alpha_{\mathrm{p}}^{-1}\alpha_{\mathrm{s}}^{-1}\alpha_{\mathrm{sa}},$$

where the morphisms are uniform, prefix, suffix and strictly alphabetic in the order of the composition. As a corollary of this result we obtain a new characterization for the rational functions as well: if $\tau$ is a rational function, then

$$\tau=\mu\alpha_{\mathrm{u}}\alpha_{\mathrm{p}}^{-1}\alpha_{\mathrm{s}}^{-1}\alpha_{\mathrm{a}},$$

where the last morphism is alphabetic.

## 2. Preliminaries

A *transducer* $T$ is a 6-tuple $(Q,\Sigma,\Delta,\delta,S,F)$ where $Q$ is a finite set of states, $\Sigma$ the input alphabet, $\Delta$ the output alphabet, $\delta$ the finite set of transitions included in $Q\times\Sigma^*\times\Delta^*\times Q$, $S\subseteq Q$ the set of starting states and $F\subseteq Q$ the set of final states.

Let the morphisms $I_T$ from $\delta^*$ into $\Sigma^*$ and $W_T$ from $\delta^*$ into $\Delta^*$ be defined by $(q,x,y,p)I_T=x$ and $(q,x,y,p)W_T=y$. A sequence $g$ of transitions $(q_i,x_i,y_i,q_{i+1})$, $i=1,2,\ldots,k$, is said to be a *computation* of $T$ from $q_1$ to $q_{k+1}$ and $g$ produces an output $gW_T=y_1y_2\ldots y_k$ from the input $gI_T=x_1x_2\ldots x_k$. For states $q,p\in Q$, we denote by $C_T(q,p)$ the set of all computations of $T$ from $q$ to $p$. By convention, the empty computation belongs to $C_T(q,q)$ for all states $q$. The set of all computations of $T$ forms a regular set $C_T$ of $\delta^*$. Further, a computation is called *accepting* if $g\in C_T(q,p)$ for some $q\in S$ and some $p\in F$. The set of all accepting computations of $T$ forms a regular set $A_T$ of $\delta^*$.

Let $T$ be a transducer defined as above. We say that $T$ *realizes* the *rational transduction* $\tau \subseteq \Sigma^* \times \Delta^*$ defined by $\tau = \{(gI_T, gW_T) \mid g \in A_T\}$ (see [4, 13]).

By $x\tau$, we denote the set $\{y \mid (x, y) \in \tau\}$. The *domain* of a rational transduction $\tau$ is the rational set $dom_\tau = \{x \in \Sigma^* \mid x\tau \neq \emptyset\}$.

The transduction $T$ is said to realize a *rational function* $\tau$ if $\tau$ is a partial function from $\Sigma^*$ into $\Delta^*$, that is, if $card(x\tau) \leqslant 1$, for all $x \in \Sigma^*$.

A transducer $T$ is said to be a *left a-gsm* if $S$ contains only one element $q_T$, $\delta$ is included in $Q \times \Sigma \times \Delta^* \times Q$ and, for all $q \in Q$ and $x \in \Sigma$, there exists at most one transition $(q, x, y, p) \in \delta$. A left a-gsm realizes a rational function.

Let $g \in C_T(q, p)$ be a computation of a left a-gsm $T$. We write $q \cdot gI_T = p$.

A rational transduction $\tau$ is said to be *left subsequential*, if there exists a left a-gsm $T_1$ realizing $\tau_1$ and a partial function $\rho$ from $F$ into $\Delta^*$ such that, for all $x \in \Sigma^*$, $x\tau = x\tau_1 (q_T \cdot x)\rho$. A left subsequential transducer $T$ is a 6-tuple $(Q, \Sigma, \Delta, \delta, q_T, \rho)$ where $(Q, \Sigma, \Delta, \delta, q_T, dom_\rho)$ is a left a-gsm and $\rho$ a partial function from $Q$ into $\Delta^*$.

The *right subsequential* transductions are defined as reversed left subsequential transductions. A rational transduction $\tau_r$ from $\Sigma^*$ into $\Delta^*$ is said to be *right subsequential*, if there exists a left subsequential transduction $\tau_l$ such that $x\tau_r = (x^r\tau_l)^r$ for all $x \in \Sigma^*$ where $x^r = x_n \ldots x_2 x_1$ is the reversal of the word $x = x_1 x_2 \ldots x_n$.

A transduction $\tau$ from $\Sigma^*$ into $\Delta^*$ is *faithful* if, for all $y \in \Delta^*$, $y\tau^{-1}$ is finite.

A transduction $\tau$ from $\Sigma^*$ into $\Delta^*$ is *positive* if $\Sigma^*\tau \subseteq \Delta^+$.

It is well known that the set of (faithful) left subsequential transductions is closed under composition [2]. The property of being positive is obviously preserved by composition.

A morphism $\alpha: \Sigma^* \to \Delta^*$ is called

- *nonerasing*, if $\Sigma\alpha \subseteq \Delta^+$,
- *alphabetic* (*strictly alphabetic*), if $\Sigma\alpha \subseteq \Delta \cup \{\varepsilon\}$ ($\Sigma\alpha \subseteq \Delta$, resp.),
- *prefix coding* (*suffix coding*), if for all $a, b \in \Sigma$ with $a \neq b$, $a\alpha$ is not a prefix (suffix, resp.) of $b\alpha$.

Moreover, if $\Sigma \cap \Delta = \emptyset$, then a morphism $\alpha: \Sigma^* \to (\Sigma \cup \Delta)^*$ is

- *uniform* (*left uniform*), if there exists a word $u \in \Delta^*$ with $a\alpha = au$ ($a\alpha = ua$, resp.) for all $a \in \Sigma$.

Let $H$ (resp. $H_a, H_{sa}, H_{ne}, H_u, H_i, H_p, H_s$) be the family of morphisms (resp. alphabetic morphisms, strictly alphabetic morphisms, nonerasing morphisms, uniform morphisms, injective morphisms, prefix codings, suffix codings). For the above sets of morphisms we let $H_x^{-1}$ denote the set of inverses $\alpha^{-1}$ of the morphisms from $H_x$.

A *right* (resp. *left*) *marker* $\mu_m$ is a mapping that sets a special symbol $m$ at the end (resp. beginning) of each word, that is, $\mu_m$ from $\Sigma^*$ into $(\Sigma \cup \{m\})^*$ is defined by $x\mu_m = xm$ (resp. $x\mu_m = mx$), for all $x \in \Sigma^*$. We denote by $M_r$ (resp. $M_l$) the family of right (resp. left) markers. A *right removal* $\nu_\Theta$ is a mapping that removes a marker from the end of the words, that is, $\nu_\Theta$ from $(\Sigma \cup \Theta)^*$ into $\Sigma^*$, where $\Sigma \cap \Theta = \emptyset$ is defined by $x\nu_\Theta = y$, if $x = ym$ with $y \in \Sigma^*$ and $m \in \Theta$, otherwise $x\nu_\Theta$ is undefined. The class of right removals is denoted by $R_r$.

Clearly compositions of markers, morphisms and inverse morphisms are rational transductions. Also the converse holds, see [9, 15].

For rational transductions the following was proved in [10, 12, 16].

**Theorem 2.1.** *The set of rational transductions is equal to*

$$M_r H^{-1} H H^{-1} H = (M_r \cup H \cup H^{-1})^* = M_r H H^{-1} H H^{-1}.$$

The next results were proved for rational functions in [7, 8].

**Theorem 2.2.** *The set of rational functions is equal to* $M_r H_i H_i^{-1} H = (M_r \cup H \cup H_i^{-1})^*$. *The set of left subsequential functions is equal to* $M_r H H_p^{-1} H = (M_r \cup H \cup H_p^{-1})^*$.

## 3. Subsequential functions

As stated in Theorem 2.2 the subsequential functions coincide with $M_r H H_p^{-1} H$, i.e., with compositions of markers, morphisms and inverses of prefix codings. We shall prove first in Lemma 3.1 that in case the functions are faithful and positive, then the morphisms can be assumed to be nonerasing. After this we strengthen this result by showing that only markers, uniform morphisms, inverses of prefix codings and strictly alphabetic morphisms are required in the compositions. On the basis of this we obtain later (Theorem 3.6) a stronger representation of the left subsequential functions.

**Lemma 3.1.** *Each positive faithful left subsequential function belongs to* $M_r H_u H_p^{-1} H_{ne}$.

**Proof.** Let $T_1 = (Q_1, \Sigma_1, \Delta_1, \delta_1, q_{T_1}, \rho_1)$ be a left subsequential transducer realizing $\tau$.

We first build a subsequential transducer $T_2 = (Q_2, \Sigma_1, \Delta_1, \delta_2, q_{T_2}, \rho_2)$ such that, for all $q \in Q_2$, $q\rho_2 \neq \emptyset$ implies $q\rho_2 \neq \varepsilon$.

For this, let us define

$$Q_2 = Q_1 \times (\{\varepsilon\} \cup \Delta_1)$$

$$q_{T_2} = (q_{T_1}, \varepsilon)$$

$$((q,\varepsilon), a, \varepsilon, (q',\varepsilon)) \in \delta_2 \quad \text{when } (q, a, \varepsilon, q') \in \delta_1$$

$$((q,b), a, \varepsilon, (q',b)) \in \delta_2 \quad \text{when } (q, a, \varepsilon, q') \in \delta_1$$

$$((q,\varepsilon), a, y, (q',c)) \in \delta_2 \quad \text{when } (q, a, yc, q') \in \delta_1$$

$$((q,b), a, by, (q',c)) \in \delta_2 \quad \text{when } (q, a, yc, q') \in \delta_1$$

$$(q,b)\rho_2 = bq\rho_1$$

$$(q,\varepsilon)\rho_2 = q\rho_1$$

The transducer $T_2$ realizes the same transduction as $T_1$ but holds up the output of one letter.

If a computation $g$ of $T_2$ ends in a state $(q, \varepsilon)$, its output $gW_{T_2}$ is equal to $\varepsilon$. Since $\Sigma^*\tau \subseteq \Delta^+$, an accepting computation of $T_2$ ends in a state $(q, b)$ with either $b \neq \varepsilon$ or $q\rho_1 \neq \varepsilon$.

Let $m$ be a new symbol which will end the input words.

We define a left a-gsm $T_3 = (Q_2, \Sigma_1 \cup \{m\}, \Delta_1, \delta_3, q_{T_2}, \{q_{T_2}\})$ by setting $\delta_3 = \delta_2 \cup \{(q, m, q\rho_2, q_{T_2}) \mid q\rho_2 \neq \emptyset\}$.

The transduction $\tau_3$ realized by $T_3$ acts on marked words: $\tau = \mu_m \tau_3$ where $\mu_m$ is a right marker.

Let $Q_2 = \{q_0, q_1, \ldots, q_n\}$ where $q_0 = q_{T_2}$ and let $d$ be a new symbol. We define a new alphabet $\Gamma$ by using the computations of $T_3$.

$$\Gamma = \{(q_{i_0}, a_1 \ldots a_r, y_r, q_{i_r}) \mid 1 \leqslant r \leqslant n,$$
$$(q_{i_0}, a_1, y_1, q_{i_1}), \ldots, (q_{i_{r-1}}, a_r, y_r, q_{i_r}) \in C_{T_3}(q_{i_0}, q_{i_r})$$
$$\text{such that } y_j = \varepsilon, \forall j < r \text{ and } y_r \neq \varepsilon\}.$$

In order to get $\tau_3 = \alpha_1 \alpha_2^{-1} \alpha_3$, let us define $\alpha_i$, $i = 1, 2, 3$, as follows:

$$a\alpha_1 = ad^n \text{ for all } a \in \Sigma_1 \cup \{m\},$$
$$(q_{i_0}, a_1 \cdots a_r, y_r, q_{i_r})\alpha_2 = d^{i_0} a_1 d^n a_2 d^n \cdots d^n a_r d^{n-i_r} \quad \text{for all } (q_{i_0}, a_1 \cdots a_r, y_r, q_{i_r}) \in \Gamma,$$
$$(q_{i_0}, a_1 \cdots a_r, y_r, q_{i_r})\alpha_3 = y_r \neq \varepsilon \quad \text{for all } (q_{i_0}, a_1 \cdots a_r, y_r, q_{i_r}) \in \Gamma.$$

The transduction $\tau$ being faithful, no computation $g$ of $T_3$ longer than $n$ has an output $gW_{T_3}$ equal to $\varepsilon$. And the last transition $(q, m, q\rho_2, q_{T_2})$ is nonerasing. This implies that any accepting computation can be factorized in nonerasing sequences of length at most $n$ ending with a nonerasing transition.

Each word $d^{i_0} a_1 d^n a_2 d^n \ldots d^n a_r d^{n-i_r}$ belonging to $\Gamma\alpha_2$ corresponds to a sequence of deterministic transitions from $q_{i_0}$ to $q_{i_r}$. The sequence will end with the first nonerasing transition. A word $d^{i_0} a_1 d^n a_2 \ldots d^n a_r d^{n-i_r}$ belonging to $\Gamma\alpha_2$ cannot be a prefix of a word $d^{i_0} a_1 d^n a_2 \ldots d^n a_s d^{n-i_s}$ in $\Gamma\alpha_2$; the transition $(q_{i_{r-1}}, a_r, y_r, q_{i_r})$ would be nonerasing in the first case and erasing in the second. Thus the morphism $\alpha_2$ maps the new alphabet $\Gamma$ into a prefix code.

By the above considerations we have then that $\tau = \mu_m \alpha_1 \alpha_2^{-1} \alpha_3 \in M_r H_u H_p^{-1} H_{ne}$. □

Right markers, uniform morphisms, nonerasing morphisms and inverses of prefix codings are faithful left subsequential functions. Preceded by a right marker, compositions of uniform morphisms, nonerasing morphisms and inverses of prefix codings are positive. Combining these observations with Lemma 3.1 leads to the following decomposition result.

**Proposition 3.2.** *The set of positive faithful left subsequential functions is equal to*

$$M_r H_u H_p^{-1} H_{ne} = M_r (H_{ne} \cup H_p^{-1})^*.$$

In the proof Lemma 3.1, if the transduction $\tau$ is not positive, we define $T_3$ directly from $T_1$ so that the transduction $\tau_3$ does not delete the marker $m$:

$$T_3=(Q_1,\Sigma_1\cup\{m\},\Delta_1,\delta_3,q_{T_1},\{q_{T_1}\}) \quad \text{where}$$

$$\delta_3=\delta_1\cup\{(q,m,q\rho_1\cdot m,q_{T_1})\mid q\rho_1\neq\emptyset\}.$$

In this way we obtain the following result:

**Proposition 3.3.** *The set of faithful left subsequential functions is equal to*

$$M_{\mathrm{r}}H_{\mathrm{u}}H_{\mathrm{p}}^{-1}H_{\mathrm{ne}}R_{\mathrm{r}}=M_{\mathrm{r}}(H_{\mathrm{ne}}\cup H_{\mathrm{p}}^{-1})^*R_{\mathrm{r}}.$$

We now demonstrate two lemmas to improve these results.

**Lemma 3.4.** $H_{\mathrm{ne}}\subset H_{\mathrm{u}}H_{\mathrm{p}}^{-1}H_{\mathrm{sa}}$ *and* $H_{\mathrm{ne}}\subset H_{\mathrm{lu}}H_{\mathrm{s}}^{-1}H_{\mathrm{sa}}$.

**Proof.** Let $\delta$ be a nonerasing morphism from $\Sigma^*$ into $\Delta^*$ where $\Sigma=\{a_i\mid 1\leqslant i\leqslant m\}$.

We define a new alphabet $\Gamma$ which will distinguish all letters in the words $a_i\delta$ for $a_i\in\Sigma$. Set $n_i=|a_i\delta|$ and $\Gamma=\{(i,j)\mid 1\leqslant i\leqslant m,\ 1\leqslant j\leqslant n_i\}$. Set $u_i=(i,1)\ldots(i,n_i)$.

Let $\alpha_k$, for $k=1,2,3$, be defined as follows:

- $\alpha_1$ from $\Sigma^*$ into $(\Sigma\cup\Gamma)^*$ such that $a_i\alpha_1=a_iu_1\cdots u_m$ for all $a_i\in\Sigma$.
- $\alpha_2$ from $\Gamma^*$ into $(\Sigma\cup\Gamma)^*$ such that

$$(i,1)\alpha_2=a_iu_1\cdots u_m \quad \text{if } n_i=1$$

$$(i,1)\alpha_2=a_iu_1\cdots u_{i-1}(i,1) \quad \text{if } n_i>1$$

$$(i,j)\alpha_2=(i,j) \quad \text{for } 1<j<n_i$$

$$(i,n_i)\alpha_2=(i,n_i)u_{i+1}\cdots u_m \quad \text{if } n_i>1$$

- $\alpha_3$ from $\Gamma^*$ into $\Delta^*$ such that $(i,j)\alpha_3=(a_i\delta)_j$ where $(a_i\delta)_j$ is the $j$th letter in the word $a_i\delta$.

Let us remark that $a_i\alpha_1=a_iu_1\cdots u_m=u_i\alpha_2$ and $u_i\alpha_3=a_i\delta$. Since the words in $\Gamma\alpha_2$ differ in their first letter, the set $\Gamma\alpha_2$ is a prefix code.

Then $\delta=\alpha_1\alpha_2^{-1}\alpha_3$ and $\delta$ belongs to $H_{\mathrm{u}}H_{\mathrm{p}}^{-1}H_{\mathrm{sa}}$.

Symmetrically, this result holds with suffix coding: $H_{\mathrm{ne}}\subset H_{\mathrm{lu}}H_{\mathrm{s}}^{-1}H_{\mathrm{sa}}$. □

In [11], it has been shown that $H_{\mathrm{ne}}^{-1}H_{\mathrm{u}}\subset H_{\mathrm{u}}H_{\mathrm{ne}}^{-1}$. However, if $H_{\mathrm{ne}}^{-1}$ is replaced by $H_{\mathrm{p}}^{-1}$, we obtain a similar result.

**Lemma 3.5.** $H_{\mathrm{p}}^{-1}H_{\mathrm{u}}\subset H_{\mathrm{u}}H_{\mathrm{p}}^{-1}$.

**Proof.** Let $\eta$ be a morphism from $\Delta^*$ into $\Sigma^*$ such that $\Delta\eta$ is a prefix code and $\gamma_u$ be a uniform morphism from $\Delta^*$ into $(\Delta\cup\Gamma)^*$ with $b\gamma_u=bu$ for all $b\in\Delta$.

Let us define the uniform morphism $\gamma_{\bar{u}}$ from $\Sigma^*$ into $(\Sigma\cup\bar{\Gamma})^*$ by, for all $a\in\Sigma$, $a\gamma_{\bar{u}}=a\bar{u}$ where $\bar{u}$ is a copy of $u$ on the alphabet $\bar{\Gamma}$ disjoint of $\Sigma$.

Since $\Delta\eta$ is a prefix code, $\Delta\eta\gamma_{\bar{u}}$ is also a prefix code. Then $\Delta\eta\gamma_{\bar{u}}\bar{u}^{-1}$ is a prefix code included in $\Sigma(\bar{u}\Sigma)^*$. The alphabets $\bar{\Gamma}$ and $\Sigma$ being disjoint, the union $\Delta\eta\gamma_{\bar{u}}\bar{u}^{-1}\cup\bar{\Gamma}$ is a prefix code.

Let us define the prefix coding $\lambda$ from $(\Delta\cup\Gamma)^*$ into $(\Sigma\cup\bar{\Gamma})^*$ by,

$$b\lambda=b\eta\gamma_{\bar{u}}\bar{u}^{-1} \quad (\forall\, b\in\Delta),$$

$$c\lambda=\bar{c} \quad (\forall\, c\in\Gamma).$$

Then $\eta^{-1}\gamma_u=\gamma_{\bar{u}}\lambda^{-1}$. □

Applying Lemma 3.4 and Lemma 3.5, we obtain the following theorem.

**Theorem 3.6.**
- *The set of positive faithful left subsequential functions is equal to* $M_rH_uH_p^{-1}H_{sa}$.
- *The set of faithful left subsequential functions is then equal to* $M_rH_uH_p^{-1}H_{sa}R_r$.
- *The set of left subsequential functions is equal to* $M_rH_uH_p^{-1}H_a$.

**Proof.** Proposition 3.2. states that the set of positive faithful left subsequential functions is equal to $M_rH_uH_p^{-1}H_{ne}$.

$$\begin{aligned} M_rH_uH_p^{-1}H_{ne} &\subseteq M_rH_uH_p^{-1}H_uH_p^{-1}H_{sa} && \text{Lemma 3.4}\\ &\subseteq M_rH_uH_uH_p^{-1}H_p^{-1}H_{sa} && \text{Lemma 3.5}\\ &\subseteq M_rH_uH_p^{-1}H_{sa} \end{aligned}$$

Then the set of positive faithful left subsequential functions is equal to $M_rH_uH_p^{-1}H_{sa}$.

Similarly, using Proposition 3.3, the set of faithful left subsequential functions is equal to $M_rH_uH_p^{-1}H_{sa}R_r$.

Since every left subsequential function is the composition of a faithful left subsequential function followed by an alphabetic morphism [2], the set of left subsequential functions is equal to $M_rH_uH_p^{-1}H_a$. □

By symmetry, we may state the same propositions concerning right subsequential functions.

**Theorem 3.7.**
- *The set of positive faithful right subsequential functions is equal to* $M_lH_{lu}H_s^{-1}H_{sa}$ *where* $H_{lu}$ *is the family of left uniform morphisms.*
- *The set of faithful right subsequential functions is equal to* $M_lH_{lu}H_s^{-1}H_{sa}R_l$.
- *The set of right subsequential functions is equal to* $M_lH_{lu}H_s^{-1}H_a$.

## 4. Rational functions

By Theorem 2.2, the set of rational functions coincides with $M_r H_i H_i^{-1} H$. We shall proceed with our study to state more precisely the set of morphisms involved in the compositions. Indeed, we shall prove that the set of rational functions coincides with $M_r H_u H_p^{-1} H_s^{-1} H_a$.

Here one should notice that $H_p^{-1} H_s^{-1} \neq H_i^{-1}$. In fact, there exists injective morphisms which are not compositions of prefix codings and suffix codings [3], e.g., the coding $\tau$ from $\{x, y, z, t\}^*$ into $\{a, b\}^*$, defined by $x\tau = a$, $y\tau = ab$, $z\tau = b^2 a$ and $t\tau = b^3 ab^4$, does not belong to $(H_p \cup H_s)^*$.

The set of positive faithful left subsequential functions being closed under composition, we easily see that $M_r H_u H_p^{-1} H_{sa} = (M_r H_u H_p^{-1} H_{sa})^* = M_r (H_u H_p^{-1} H_{sa})^*$ and, by Proposition 3.2, it is also equal to $M_r (H_{ne} \cup H_p^{-1})^*$.

We shall first demonstrate that these equalities hold also without the right markers, i.e. $(H_{ne} \cup H_p^{-1})^* = H_u H_p^{-1} H_{sa}$.

**Lemma 4.1.** $H_{sa} H_u \subset H_u H_{sa}$.

**Proof.** Let $\alpha$ be a strictly alphabetic morphism from $\Sigma^*$ into $\Delta^*$ and $\gamma_u$ be a uniform morphism from $\Delta^*$ into $(\Delta \cup \Gamma)^*$ with $b\gamma_u = bu$ for all $b \in \Delta$.

Let us define the uniform morphism $\gamma_{\bar{u}}$ from $\Sigma^*$ into $(\Sigma \cup \bar{\Gamma})^*$ by, for all $a \in \Sigma$, $a\gamma_{\bar{u}} = a\bar{u}$ where $\bar{u}$ is a copy of $u$ on the alphabet $\bar{\Gamma}$ disjoint of $\Sigma$.

Let us define the strictly alphabetic morphism $\beta$ from $(\Sigma \cup \bar{\Gamma})^*$ into $(\Delta \cup \Gamma)^*$ by, for all $a \in \Sigma$, $a\beta = a\alpha$ and, for all $\bar{c} \in \bar{\Gamma}$, $\bar{c}\beta = c$.

Then $\alpha\gamma_u = \gamma_{\bar{u}}\beta$. □

**Lemma 4.2.** $H_{sa} H_p^{-1} \subset H_p^{-1} H_{sa}$ *and* $H_{sa} H_s^{-1} \subset H_s^{-1} H_{sa}$.

**Proof.** Let $\alpha$ be a strictly alphabetic morphism from $\Sigma^*$ into $\Delta^*$ and $\eta$ be a morphism from $\Gamma^*$ into $\Delta^*$ such that $\Gamma\eta$ is a prefix code.

$\Gamma\eta\alpha^{-1}$ is also a prefix code. If there exists $x, y \in \Gamma\eta\alpha^{-1}$ such that $x$ is a left factor of $y$, then $x\alpha$ would be a left factor of $y\alpha$ but $x\alpha, y\alpha \in \Gamma\eta$.

Let us define the alphabet $\Omega = \{c_w \mid w \in \Gamma\eta\alpha^{-1}\}$.

Let us define the prefix coding $\lambda$ from $\Omega^*$ into $\Sigma^*$ by, for all $c_w \in \Omega$, $c_w\lambda = w$ such that $w \in c\eta\alpha^{-1}$.

Let us define the strictly alphabetic morphism $\beta$ from $\Omega^*$ into $\Gamma^*$ by, for all $c_w \in \Omega$, $c_w\beta = c$ such that $c_w\lambda \in c\eta\alpha^{-1}$.

Then $\alpha\eta^{-1} = \lambda^{-1}\beta$. Thus $H_{sa} H_p^{-1} \subset H_p^{-1} H_{sa}$.

A similar reasoning leads to $H_{sa} H_s^{-1} \subset H_s^{-1} H_{sa}$. □

**Theorem 4.3.** $(H_{ne} \cup H_p^{-1})^* = (H_u H_p^{-1} H_{sa})^* = H_u H_p^{-1} H_{sa}$.

**Proof.** By Lemma 3.4, we have that $H_{\mathrm{ne}} \subset H_{\mathrm{u}} H_{\mathrm{p}}^{-1} H_{\mathrm{sa}}$. Also,

$$
\begin{aligned}
H_{\mathrm{u}} H_{\mathrm{p}}^{-1} H_{\mathrm{sa}} H_{\mathrm{u}} H_{\mathrm{p}}^{-1} H_{\mathrm{sa}} &\subseteq H_{\mathrm{u}} H_{\mathrm{p}}^{-1} H_{\mathrm{u}} H_{\mathrm{sa}} H_{\mathrm{p}}^{-1} H_{\mathrm{sa}} && \text{Lemma 4.1} \\
&\subseteq H_{\mathrm{u}} H_{\mathrm{p}}^{-1} H_{\mathrm{u}} H_{\mathrm{p}}^{-1} H_{\mathrm{sa}} H_{\mathrm{sa}} && \text{Lemma 4.2} \\
&\subseteq H_{\mathrm{u}} H_{\mathrm{u}} H_{\mathrm{p}}^{-1} H_{\mathrm{p}}^{-1} H_{\mathrm{sa}} H_{\mathrm{sa}} && \text{Lemma 3.5} \\
&\subseteq H_{\mathrm{u}} H_{\mathrm{p}}^{-1} H_{\mathrm{sa}}. \quad \square
\end{aligned}
$$

We refer to [1, 4, 5, 14] for the definitions of left (right) sequential functions and for the following decomposition result, which links rational functions and sequential functions together.

**Theorem 4.4.** *Let $\tau$ be a rational function from $\Sigma^*$ into $\Delta^*$ such that $\varepsilon\tau = \varepsilon$. There exist a right (left) length preserving sequential function $\sigma_1$ from $\Sigma^*$ into $\Gamma^*$ and a left (right) sequential function $\sigma_2$ from $\Gamma^*$ into $\Delta^*$ such that $\tau = \sigma_1\sigma_2$.*

The fact that the first sequential function $\sigma_1$ can be assumed to be length preserving follows from the proof of the decomposition theorem as given, e.g., in [1].

We shall now prove this decomposition result for positive and faithful rational functions $\tau$. Notice that such a $\tau$ does not satisfy the property $\varepsilon\tau = \varepsilon$.

**Theorem 4.5.** *Let $\tau$ be a positive faithful rational function from $\Sigma^*$ into $\Delta^*$. There exist a positive faithful right (left) subsequential function $\sigma_1$ from $\Sigma^*$ into $\Gamma^*$ and a positive faithful left (right) subsequential function $\sigma_2$ from $\Gamma^*$ into $\Delta^*$ such that $\tau = \sigma_1\sigma_2$.*

**Proof.** Consider a positive faithful rational function $\tau$.

Let $a_\varepsilon$ be a new letter, and let us define a positive and faithful left (and right) subsequential function $\sigma_0$ by

$$\varepsilon\sigma_0 = a_\varepsilon \quad \text{and} \quad v\sigma_0 = v \quad (\forall\, v \in \Sigma^+).$$

It is immediate that $\tau = \sigma_0\tau_1$, where $\varepsilon\tau_1 = \varepsilon$, $a_\varepsilon\tau_1 = \varepsilon\tau$ and $v\tau_1 = v\tau$ for all nonempty $v \in \Sigma^+$.

By Theorem 4.4, there are right sequential function $\sigma_1$ and a left sequential function $\sigma_2$ such that $\tau_1 = \sigma_1\sigma_2$, where $\sigma_1$ is length preserving and hence faithful. However, $\varepsilon\sigma_1 = \varepsilon$ and thus $\sigma_1$ is not positive.

The function $\tau_1$ is faithful and $\sigma_1$ is length preserving imply together that the second sequential function $\sigma_2$ is also faithful. Again, it need not be positive.

We overcome the problem of positiveness by defining, for $i = 1, 2$, $\sigma'_i$ by

$$\varepsilon\sigma'_i = \emptyset \quad \text{and} \quad v\sigma'_i = v\sigma_i \quad (\forall\, v \neq \varepsilon).$$

Now, let $\tau'_1 = \sigma'_1\sigma'_2$. We have then that $\tau = \sigma_0\sigma'_1\sigma'_2$, since

$$\varepsilon\sigma_0\sigma'_1\sigma'_2 = a_\varepsilon\sigma'_1\sigma'_2 = a_\varepsilon\sigma_1\sigma'_2 = a_\varepsilon\sigma_1\sigma_2 = a_\varepsilon\tau_1 = \varepsilon\tau$$

and, clearly, $v\sigma_0\sigma'_1\sigma'_2 = v\tau$ for all $v \in \Sigma^+$.

The claim follows now since $\tau=(\sigma_0\sigma'_1)\sigma'_2$ is a required composition of $\tau$ into positive and faithful right and left subsequential functions. □

**Theorem 4.6.** *The set of positive faithful rational functions is equal to*

$$M_r H_u H_p^{-1} H_s^{-1} H_{sa}.$$

**Proof.** By Theorem 4.5, $\tau=\sigma_1\sigma_2$ where $\sigma_1$ is a positive faithful left subsequential function and $\sigma_2$ a positive faithful right subsequential function.

The left subsequential function $\sigma_2$ thus belongs to $M_l H_{lu} H_s^{-1} H_{sa}$. A left marker followed by a left uniform morphism is a positive faithful left subsequential function. Thus, $M_r H_u H_p^{-1} H_{sa} M_l H_{lu}$ is included into $M_r H_u H_p^{-1} H_{sa}$.

$$\begin{aligned} \tau &\subseteq M_r H_u H_p^{-1} H_{sa} H_s^{-1} H_{sa} \\ &\subseteq M_r H_u H_p^{-1} H_s^{-1} H_{sa} H_{sa} \quad \text{Lemma 4.3} \\ &\subseteq M_r H_u H_p^{-1} H_s^{-1} H_{sa} \end{aligned}$$

Markers, morphisms and inverse of codings are faithful rational functions which are closed under composition. The marker makes the composition positive. □

By symmetry, we may state that the set of positive faithful rational functions is equal to $M_l H_{lu} H_s^{-1} H_p^{-1} H_{sa}$.

A faithful function is the composition of a positive faithful function followed by a removal. Then, the set of faithful rational functions is equal to $M_r H_u H_p^{-1} H_s^{-1} H_{sa} R_r = M_l H_{lu} H_s^{-1} H_p^{-1} H_{sa} R_l$.

Since a rational function is a faithful rational function followed by an alphabetic morphism and since prefix codings or suffix codings are injective morphisms, we have the following characterization for rational functions.

**Corollary 4.7.** *The set of rational functions is equal to*

$$M_r H_u H_p^{-1} H_s^{-1} H_a = M_l H_{lu} H_s^{-1} H_p^{-1} H_a = M_r H_u H_i^{-1} H_a.$$

## References

[1] A. Arnold and M. Latteux, A new proof of two theorems about rational transductions, *Theoret. Comput. Sci.* **8** (1979) 261–263.

[2] J. Berstel, *Transductions and Context-Free Languages* (Teubner, Stuttgart, 1979).

[3] J. Berstel and D. Perrin, *Theory of Codes* (Academic Press, New York, 1985).

[4] S. Eilenberg, *Automata, Languages and Machines, Vol A* (Academic Press, New York, 1974).

[5] C.C. Elgot and J.E. Mezei, On relations defined by generalized finite automata, *IBM J. Res. Develop.* **9** (1965) 47–68.

[6] T. Harju and H.C.M. Kleijn, Decidability problems for unary output sequential transducers, *Discrete Appl. Math.* **32** (1991) 131–140.

[7] T. Harju, H.C.M. Kleijn and M. Latteux, Compositional representation of rational functions, *RAIRO Théor. Inform Appl.* **26** (1992) 243–255.
[8] T. Harju, H.C.M. Kleijn and M. Latteux, Deterministic rational functions, *Acta Inform.* **29** (1992) 545–554.
[9] J. Karhumäki and M. Linna, A note on morphic characterization of languages, *Discrete Appl. Math.* **5** (1983) 243–246.
[10] M. Latteux and J. Leguy, On the composition of morphisms and inverse morphisms, Lecture Notes in Computer Science, Vol. 154 (Springer, Berlin, 1983) 420–432.
[11] M. Latteux and E. Timmerman, Bifaithful starry transductions, *Inform. Process. Lett.* **28** (1988) 1–4.
[12] M. Latteux and P. Turakainen, A new normal form for the composition of morphisms and inverse morphisms, *Math. Systems Theory* **20** (1987) 261–271.
[13] M. Nivat, Transductions des langages de Chomsky, *Ann. Inst. Fourier* **18** (1968) 339–456.
[14] M.P. Schützenberger, Sur les relations rationnelles entre monoïdes libres., *Theoret. Comput. Sci.* **3** (1976) 243–259.
[15] P. Turakainen, A homomorphic characterization of principal semi-AFLs without using intersection with regular sets, *Inform. Sci.* **27** (1982) 141–149.
[16] P. Turakainen, A machine-oriented approach to composition of morphisms and inverse morphisms, *EATCS Bull.* **20** (1983) 162–166.